

明細書

監視装置

技術分野

5

本発明は、監視装置に関し、特にデジタル・ビデオ／オーディオ信号の監視に適した監視装置に関する。

背景技術

10

近年の映像処理技術の向上により、ハイビジョンテレビ放送など高画質な映像が放映されるようになってきている。ここで、ハイビジョン放送などにかかるデジタル映像信号は、衛星放送やケーブルTVのネットワークを介して各家庭に伝送されることが多い。しかるに、映像信号が伝送される間に、種々の原因によりエラーが生じることがある。エラーが生じると、映像のフリーズ、ブラックアウト、ノイズ、音声ミュートなどの不具合を招く恐れあり、その対策が必要となっている。

15

これに対し本出願人は、特開2003-204562において、例えば中央処理端末が、伝送元から出力された映像信号（第1の信号）に基づく第1の統計値と、中継局又は伝送先から出力された映像信号（第2の信号）に基づく第2の統計値との差をとり、その差が、閾値を下回っていれば正常と判断し、閾値を上回っていれば、伝送元10と中継局20との間で伝送異常が生じたと判断して、警報信号を出力し警報（警報表示や警報音）を発する信号監視システムを開示する。

25

しかるに、かかる信号監視システムの場合には、中央端末が、第1の統計値と第2の統計値との差をとって、その差に基づいて自動的に伝送異常を判別するのみであるため、それだけでは、いかなるエラーが生じたか十分な解析を行えないという問題がある。エラーの解析を十分に行えないと、また同様なトラブルを招く恐れがある。ここで、エラーを解析するためには、送信された映像信号を全て記憶した上で、後から時間をかけて再生しながらチェックする必要があるが、そのためには映像音声信号を記憶する膨大な記憶容量と、膨大なチェック時間が必要とされる。

10 発明の開示

本発明は、かかる従来技術の問題点に鑑みてなされたものであり、エラーが生じた場合に、エラー内容を解析できる監視システムを提供することを目的とする。

15

本発明の監視システムは、伝送元から伝送先へと伝送される映像音声信号を監視する監視システムにおいて、

伝送元から伝送先へと伝送された映像音声信号を、所定時間だけ繰り返し記憶するステップと、

20 伝送される前の映像音声信号から抽出される第1の特徴量と、伝送された後の映像音声信号から抽出される第2の特徴量とをリアルタイムで比較するステップと、

前記第1の特徴量と前記第2の特徴量との間に所定値以上の差がある場合、エラーが生じたと判定するステップと、

25 エラーが生じたと判定されたときは、記憶された映像音声信号を所定の宛先に送信するステップとを有することを特徴とする。

本発明の監視システムによれば、伝送元から伝送先へと伝送された映像音声信号を、所定時間だけ繰り返し記憶するので、例えば所定のタイミングで映像音声信号を上書きするようにすれば、映像音声信号の記憶容量が少なくて済む。

5 またエラーが生じたと判定された時点では、エラーの生じた映像音声信号は、上書き等されることなく残存しているので、これを所定の宛先に送信することによって、送信された映像音声信号に基づいて、エラーの解析を迅速且つ詳細に行うことができる。なお、「映像音声信号」とは、一般的には映像信号（ビデオ信号）と音声信号（オーディオ信号）の双方を含むものをいうが、本明細書

10 中では少なくとも一方を含む信号であれば足り、生データ、圧縮データのいずれも問わない。

比較のために用いられる前記第2の特徴量と、前記記憶された映像音声信号は、インターネットを介して伝送先から伝送元に送信されると好ましい。

15

前記記憶された映像音声信号は、エラーを解析するために用いられると好ましい。

20

前記エラーは、画像フリーズ現象であると好ましい。

前記エラーは、ブラックアウト現象であると好ましい。

前記エラーは、音声ミュート現象であると好ましい。

25

前記エラーは、音声不良現象であると好ましい。

前記エラーは、映像音声不整合現象であると好ましい。

前記エラーは、不正フレーム現象であると好ましい。

- 5 前記第1の特徴量と前記第2の特徴量との間に、所定値以上の差がある場合、
伝送先に伝送された映像音声信号を補正すると好ましい。

図面の簡単な説明

- 10 図1は、本実施の形態にかかる監視システムを含む伝送システム全体の概念
図である。

図2は、抽出装置100X、100A、100Bの構成を示すブロック図で
ある。

15

図3Aは、抽出装置100Xで抽出された特徴量を示す図である。図3Cは、
抽出装置100A又は100Bで抽出された特徴量を示す図である。図3Bは、
2つの特徴量の差分をとった値を示す図である。

- 20 図4Aは、抽出装置100Xで抽出された特徴量を示す図である。図4Cは、
抽出装置100A又は100Bで抽出された特徴量を示す図である。図4Bは、
2つの特徴量の差分をとった値を示す図である。

- 25 図5Aは、抽出装置100Xで抽出された特徴量を示す図である。図5Cは、
抽出装置100A又は100Bで抽出された特徴量を示す図である。図5Bは、
2つの特徴量の差分をとった値を示す図である。

図6 Aは、抽出装置100 Xで抽出された特徴量を示す図である。図6 Cは、抽出装置100 A又は100 Bで抽出された特徴量を示す図である。図6 Bは、2つの特徴量の差分をとった値を示す図である。

5

図7 Aは、抽出装置100 Xで抽出された特徴量を示す図である。図7 Cは、抽出装置100 A又は100 Bで抽出された特徴量を示す図である。図7 Bは、2つの特徴量の差分をとった値を示す図である。

10

図8 Aは、抽出装置100 Xで抽出された特徴量を示す図である。図8 Cは、抽出装置100 A又は100 Bで抽出された特徴量を示す図である。図8 Bは、2つの特徴量の差分をとった値を示す図である。

発明を実施するための最良の形態

15

以下、実施の形態を参照して本発明を説明する。図1は、本実施の形態にかかる監視システムを含む伝送システム全体の概念図である。図1において、例えば放送局などの伝送元10から、サテライト局などの伝送先20 A、20 Bに、オーディオ信号とビデオ信号とを含む映像音声信号が伝送されている場合を20
考える。かかる映像音声信号の伝送は、通信衛星Sを介して行う例を示すが、光ファイバーなど様々な形態により伝送されて良い。

25

図2は、抽出装置100 X、100 A、100 Bの構成を示すブロック図である。映像音声信号のうち左右のオーディオ信号AL、ARは、オーディオ入力部101、102に入力され、そこから出力された信号は、各々ディレイ部103、104に入力され、オーディオ演算部105で演算された結果が、音

声の特徴量 (Audio Level、Audio Activity) として出力され、抽出装置100X、100A、100Bから端末200X、200A、200Bに出力される。ここで、Audio Level とは、画像の1フレーム（例えば30フレーム/秒）に含まれる音声サンプリング（48KHz）の値（48000/30=1600個）の絶対値の平均値の値をさす。又、Audio Activity とは、画像の1フレーム（例えば30フレーム/秒）に含まれる音声サンプリング（48KHz）の値（48000/30=1600個）の2乗平均値の値をさす。

一方、映像音声信号のうちビデオ信号VDは、ビデオ入力部108に入力され、そこから出力された信号は、フレームメモリ109、110、111に入力される。フレームメモリ109は現在のフレームを記憶し、フレームメモリ110は、一つ前のフレームを記憶し、フレームメモリ111は、2つ前のフレームを記憶する。

15

フレームメモリ109、110、111からの出力信号は、MC演算部112に入力され、その演算結果が映像の特徴量 (Motion) として出力される。一方、フレームメモリ110からの出力信号は、ビデオ演算部119に入力される。ビデオ演算部119の演算結果は、映像の特徴量 (Video Level、Video Activity) として出力される。これらの出力信号は、映像の特徴量として、抽出装置100X、100A、100Bから端末200X、200A、200Bに出力される。ここで、Motion とは、画像フレームを例えば、8画素×8ラインサイズの小ブロックに分けて、この小ブロックごとに、64画素の平均値と分散を求め、Nフレーム前の同じ場所のブロックの平均値と分散値との差で表され、画像の動きを示すものである。但し、Nは通常、1、2、4のいずれかである。又、Video Level とは、画像フレーム

25

に含まれる画素の値の平均値である。更に、Video Activity としては、画像に含まれる小ブロックごとに分散を求めたとき、この分散のフレーム内の画素の平均値を用いても良いし、単純に画像フレームに含まれる画素のフレーム内での分散値を用いても良い。

5

次に、本監視システムの動作について説明する。伝送元 10 から伝送される前の映像音声信号は、それから特徴量（メタデータ）を抽出する抽出装置 100 X に入力されると共に、ビデオサーバー 201 X に所定時間だけ上書きされながら記憶される。

10

一方、伝送先 20 A、20 B に伝送された前の映像音声信号は、それから特徴量を抽出する抽出装置 100 A、100 B にそれぞれ入力されると共に、ビデオサーバー 201 A、201 B に所定時間だけ上書きされながら記憶される

（伝送元から伝送先へと伝送された映像音声信号を、所定時間だけ繰り返し記憶するステップ）。抽出された特徴量は、端末 200 A、200 B からインターネット INT を介して、伝送元 10 の端末 200 X に伝達される。

15

伝送元 10 の端末 200 X は、抽出装置 100 X が抽出した特徴量と、端末 200 A、200 B から伝達された特徴量とをリアルタイムで比較し（伝送される前の映像音声信号から抽出される第 1 の特徴量と、伝送された後の映像音声信号から抽出される第 2 の特徴量とをリアルタイムで比較するステップ）、所定値以上の差がある場合には、エラーが生じたと判断し（前記第 1 の特徴量と前記第 2 の特徴量との間に所定値以上の差がある場合、エラーが生じたと判定するステップ）、インターネット INT を介して端末 200 A、200 B に指令を送信する。端末 200 A、200 B は、かかる指令に応じて、ビデオサーバー 201 A、201 B より現時点から過去 10 秒間の映像音声信号を切り取っ

20

25

て、そのビデオクリップをインターネット INT を介して伝送元 10 の端末 200X に送信する（エラーが生じたと判定されたときは、記憶された映像音声信号を所定の宛先に送信するステップ）。

- 5 伝送元 10 の作業者は、端末 200A、200B から端末 200X に送信された映像音声信号と、ビデオサーバー 201X に記憶されていた映像音声信号とを比較して、エラーが生じた原因を解析することができる。なお、端末 200X からインターネット INT を介して、端末 200A、200B に、伝送前の映像音声信号にかかる特徴量を伝達し、伝送先で特徴量の比較を行っても良い。
- 10 い。

次に、より具体的なエラーについて説明する。

（１）画像フリーズ現象の検出

- 図 3A は、抽出装置 100X で抽出された特徴量を示す図であり、図 3C は、抽出装置 100A 又は 100B で抽出された特徴量を示す図であり、図 3B は、2 つの特徴量の差分をとった値を示す図であり、縦軸は特徴量としての motion [値] を表し、横軸は時間を表す。
- 15

- ここで、図 3C に示すように、伝送された後の映像音声信号に基づく映像においては、時間 $t_1 \sim t_2$ の間は、motion [値] が低い。図 3A に示すように、伝送される前の映像音声信号に基づく映像においても、時間 $t_1 \sim t_2$ の間は、motion [値] が低くなっており、その差分はゼロである（図 3B 参照）。これは、伝送した映像が静止画であるために生じたものであり、従って画像フリーズ現象は生じていないと判断できる。
- 20

25

一方、図 3C に示すように、伝送された後の映像音声信号に基づく映像にお

いては、時間 $t_3 \sim t_4$ の間は、motion [値] が低いのに対し、図 3 A に示すように、伝送される前の映像音声信号に基づく映像においては、時間 $t_3 \sim t_4$ の間は、motion [値] が高くなっており、その差分は閾値 TH_1 を超えている (図 3 B 参照)。これは、伝送された映像において、何らかの原因により画像
5 フリーズ現象が生じたことによるものであり、これを検出した伝送元 10 の端末 200 X は、直ちにビデオクリップを送信するように、不具合が生じた伝送先の端末 200 A 又は端末 200 B に指令を送信する。

(2) ブラックアウト現象の検出

10 図 4 A は、抽出装置 100 X で抽出された特徴量を示す図であり、図 4 C は、抽出装置 100 A 又は 100 B で抽出された特徴量を示す図であり、図 4 B は、2 つの特徴量の差分をとった値を示す図であり、縦軸は特徴量としての Video Activity [値] を表し、横軸は時間を表す。この Video Activity [値] としては、例えば以下の分散 A を用いることができる。

15

伝送前後におけるビデオ信号 (バーチャルビデオ信号のような 3 次元座標値毎に値を持つ例とする、但し $z = 0$ とすれば通常の 2 次元ビデオ信号になる) を考えたとき、時刻 t における 3 次元座標 (x, y, z) における伝送前のビデオ信号を $V(x, y, z, t)$ とし、時刻 t における 3 次元座標 (x, y, z) における伝送後のビデオ信号を $U(x, y, z, t)$ とする。
20

ここで、ビデオ信号を長い距離にわたって伝送すると、信号の欠損、ノイズなど様々な問題が生じる恐れがあるため、必ずしも $V(x, y, z, t) = U(x, y, z, t)$ とはならないが、視聴者が気づかない程度のエラーであれば補正する必要はないといえる。しかしながら、ブラックアウト現象のような
25 不具合であれば、対策が必要である。

ビデオ信号 $V(x, y, z, t)$ の特徴量としての分散 A は、以下の式で表すことができる。

$$\text{分散 } A = \frac{1}{XYZT} \sum_{x=1}^X \sum_{y=1}^Y \sum_{z=1}^Z \sum_{t=1}^T \{V(x, y, z, t) - \text{ave. } V\}^2$$

5

又、平均値 $\text{ave. } V$ は、以下の式で得ることができる。

$$\text{平均値 ave. } V = \frac{1}{XYZT} \sum_{x=1}^X \sum_{y=1}^Y \sum_{z=1}^Z \sum_{t=1}^T V(x, y, z, t)$$

10 伝送前のビデオ信号 $V(x, y, z, t)$ と、伝送後のビデオ信号を $U(x, y, z, t)$ とについて、それぞれ分散 A を求めた上で、その差分を取ること
で、以下のようにしてブラックアウトを判別できる。

図 4 C に示すように、伝送された後の映像音声信号に基づく映像においては、
時間 $t_1 \sim t_2$ の間は、分散値が低い。図 4 A に示すように、伝送される前
15 の映像音声信号に基づく映像においても、時間 $t_1 \sim t_2$ の間は、分散値が低
くなっており、その差分はゼロである (図 4 B 参照)。これは、伝送した映像が

例えば星空等を映したものであるために生じたものであり、従って画像フリーズ現象は生じていないと判断できる。

一方、図4Cに示すように、伝送された後の映像音声信号に基づく映像においては、時間 $t_3 \sim t_4$ の間は、分散値が低いのに対し、図4Aに示すように、伝送される前の映像音声信号に基づく映像においては、時間 $t_3 \sim t_4$ の間は、分散値が高くなっており、その差分は閾値 TH_2 を超えている（図4B参照）。これは、伝送された映像において、何らかの原因により画面が真っ黒になるブラックアウト現象が生じたことによるものであり、これを検出した伝送元10の端末200Xは、直ちにビデオクリップを送信するように、不具合が生じた伝送先の端末200A又は端末200Bに指令を送信する。

(3) 音声ミュート現象の検出

図5Aは、抽出装置100Xで抽出された特徴量を示す図であり、図5Cは、抽出装置100A又は100Bで抽出された特徴量を示す図であり、図5Bは、2つの特徴量の差分をとった値を示す図であり、縦軸は特徴量としての Audio Level[値]を表し、横軸は時間を表す。なお、オーディオ信号における Audio Level[値]のサンプリングは、ビデオ信号におけるフレームの周波数で平均化されると好ましい。例えば、1秒間に30フレームのビデオ信号の場合、Audio Level[値]のサンプリングは30Hzで行うと好ましい。

ここで、図5Cに示すように、伝送された後の映像音声信号に基づく音声においては、時間 $t_1 \sim t_2$ の間は、Audio Level[値]が非常に低い。図5Aに示すように、伝送される前の映像音声信号に基づく音声においても、時間 $t_1 \sim t_2$ の間は、Audio Level[値]が低くなっており、その差分はゼロである（図5B参照）。これは、伝送される前の映像音声信号において、元々の Audio

Level [値] が低かったものであり、従って音声ミュート現象は生じていないと判断できる。

一方、図5 Cに示すように、伝送された後の映像音声信号に基づく音声においては、時間 $t_3 \sim t_4$ の間は、Audio Level [値] が低いのに対し、図5 Aに示すように、伝送される前の映像音声信号に基づく音声においては、時間 $t_3 \sim t_4$ の間は、Audio Level [値] が高くなっており、その差分は閾値 TH_3 を超えている（図5 B参照）。これは、伝送された音声において、何らかの原因により音声途切れる音声ミュート現象が生じたことによるものであり、これを検出した伝送元10の端末200 Xは、直ちにビデオクリップを送信するように、不具合が生じた伝送先の端末200 A又は端末200 Bに指令を送信する。

(4) 音声不良現象の検出

図6 Aは、抽出装置100 Xで抽出された特徴量を示す図であり、図6 Cは、抽出装置100 A又は100 Bで抽出された特徴量を示す図であり、図6 Bは、2つの特徴量の差分をとった値を示す図であり、縦軸は特徴量としての Audio Level [値] を表し、横軸は時間を表す。なお、オーディオ信号における Audio Level [値] のサンプリングは、ビデオ信号におけるフレームの周波数で平均化されると好ましい。

20

ここで、伝送された後の映像音声信号に基づく Audio Level [値] と、伝送される前の映像音声信号に基づく Audio Level [値] との差分を取ったとき、図6 Bに示すように、時間 $t_1 \sim t_2$ の間、及び時間 $t_3 \sim t_4$ の間は、差分が閾値 TH_4 を超えている。これは、伝送された音声において、何らかの原因によりノイズ等が重畳され、音声不良現象が生じたことによるものであり、これを検出した伝送元10の端末200 Xは、直ちにビデオクリップを送信するよう

25

に、不具合が生じた伝送先の端末 200A 又は端末 200B に指令を送信する。

(5) 映像音声不整合現象の検出

図 7A は、ビデオフレームに対応して、抽出装置 100X で抽出された Audio Level [値] を示す図である。図 7C は、抽出装置 100A 又は 100B で抽出された Audio Level [値] を示す図である。図 7B は、時間に対して音声の進み／遅れを示す図である。なお、オーディオ信号における Audio Level [値] のサンプリングは、ビデオ信号におけるフレームの周波数で平均化されると好ましい。

10

ここでは、フレームに対する Audio Level [値] の立ち上がりを検出して比較する。図 7B に示すように、時間 t_1 、 t_3 においては、映像に対する音声の遅れ量が閾値 TH_5+ を超えており、時間 t_2 においては、映像に対する音声の進み量が閾値 TH_5- を下回っている。いずれかを検出することで、映像音声不整合現象が生じたと判断した伝送元 10 の端末 200X は、直ちにビデオクリップを送信するように、不具合が生じた伝送先の端末 200A 又は端末 200B に指令を送信する。

15

(6) 不正フレーム現象の検出

図 8A は、抽出装置 100X で抽出された特徴量を示す図であり、図 8C は、抽出装置 100A 又は 100B で抽出された特徴量を示す図であり、図 8B は、2 つの特徴量の差分をとった値を示す図であり、縦軸は特徴量としての Video Activity [値] (上述した分散を用いて良い) を表し、横軸は時間を表す。

20

伝送される前の映像音声信号に基づく映像と、伝送された後の映像音声信号に基づく映像とが完全に一致している場合、画像値の統計量の差分を取るとゼ

25

口になる。ところが、伝送された後の映像音声信号中に、1フレームだけ異なる映像の信号が挿入されたような場合、その差分が所定の閾値を超える。

図8Bに示すように、時間 $t_1 \sim t_2$ の間において、画素値の統計量の差分が閾値 TH_6+ を超えており、また時間 $t_3 \sim t_4$ の間においては、画素値の統計量の差分が閾値 TH_6- を下回っている。いずれかを検出することで、不正フレーム現象が生じたと判断した伝送元10の端末200Xは、直ちにビデオクリップを送信するように、不具合が生じた伝送先の端末200A又は端末200Bに指令を送信する。

10

なお、伝送元10においてエラーの解析が行われた後、例えば、映像音声不整合現象が生じた場合には、ビデオ信号とオーディオ信号とのズレをなくすような指令を、伝送元の端末200Xから、伝送先の端末200A又は200Bへと送信し、それに基づいて伝送先の端末200A又は200Bが、映像音声信号において、ビデオ信号とオーディオ信号とのズレをなくす処理を行うことができる。かかる指令は、インターネットを用いることで、伝送先の端末に一括して送信することができる。

15

請求の範囲

(1) 伝送元から伝送先へと伝送される映像音声信号を監視する監視システムにおいて、

5 伝送元から伝送先へと伝送された映像音声信号を、所定時間だけ繰り返し記憶するステップと、

伝送される前の映像音声信号から抽出される第1の特徴量と、伝送された後の映像音声信号から抽出される第2の特徴量とをリアルタイムで比較するステップと、

10 前記第1の特徴量と前記第2の特徴量との間に所定値以上の差がある場合、エラーが生じたと判定するステップと、

エラーが生じたと判定されたときは、記憶された映像音声信号を所定の宛先に送信するステップとを有することを特徴とする監視システム。

15 (2) 比較のために用いられる前記第2の特徴量と、前記記憶された映像音声信号は、インターネットを介して伝送先から伝送元に送信されることを特徴とする請求項1に記載の監視システム。

20 (3) 前記記憶された映像音声信号は、エラーを解析するために用いられることを特徴とする請求項1又は2に記載の監視システム。

(4) 前記エラーは、画像フリーズ現象であることを特徴とする請求項1～3のいずれかに記載の監視システム。

25 (5) 前記エラーは、ブラックアウト現象であることを特徴とする請求項1～4のいずれかに記載の監視システム。

(6) 前記エラーは、音声ミュート現象であることを特徴とする請求項 1～5 のいずれかに記載の監視システム。

5 (7) 前記エラーは、音声不良現象であることを特徴とする請求項 1～6 のいずれかに記載の監視システム。

(8) 前記エラーは、映像音声不整合現象であることを特徴とする請求項 1～7 のいずれかに記載の監視システム。

10

(9) 前記エラーは、不正フレーム現象であることを特徴とする請求項 1～8 のいずれかに記載の監視システム。

15 (10) 前記第 1 の特徴量と前記第 2 の特徴量との間に、所定値以上の差がある場合、伝送先に伝送された映像音声信号を補正することを特徴とする請求項 1～9 のいずれかに記載の監視システム。

1/8

FIG. 1

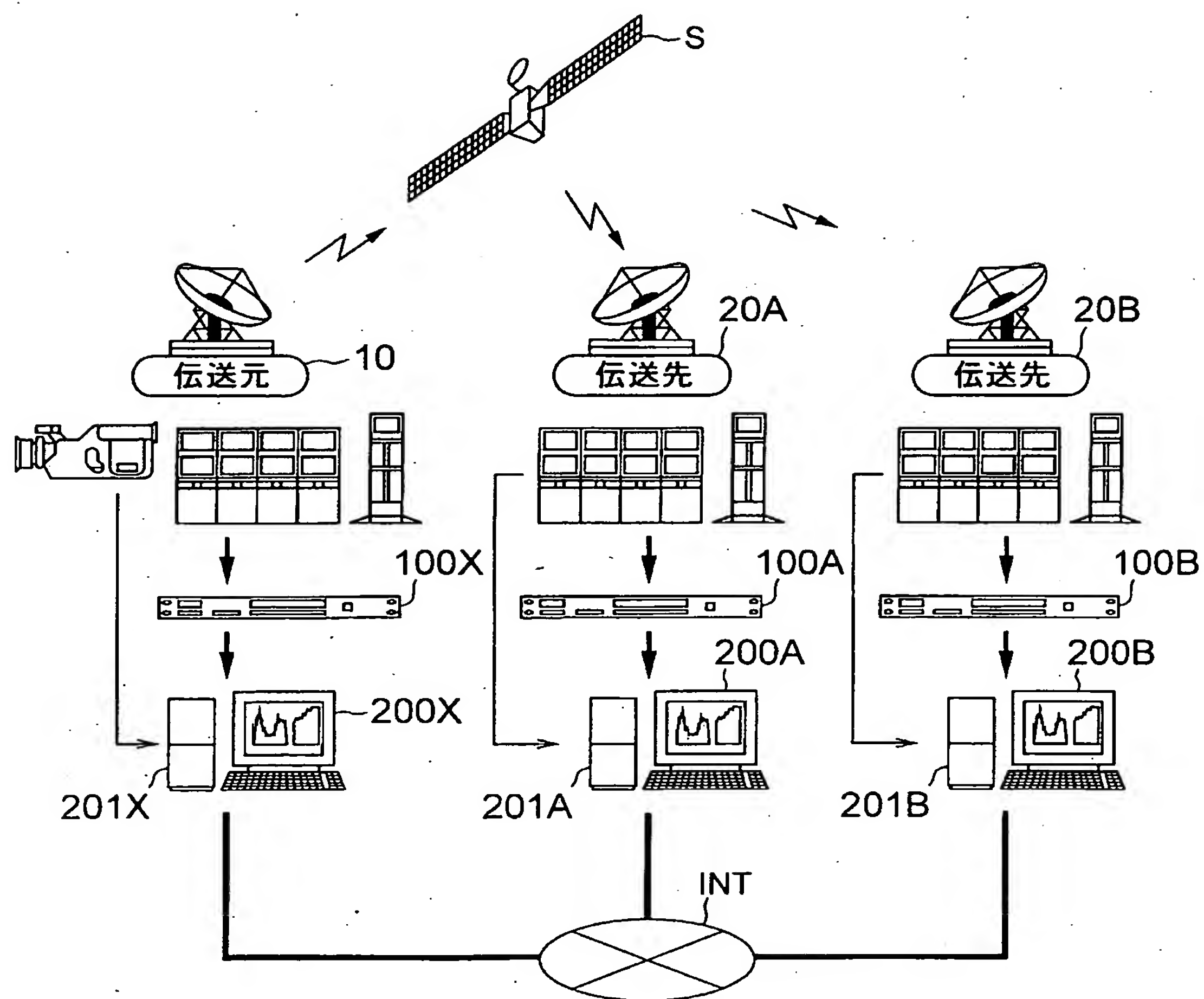
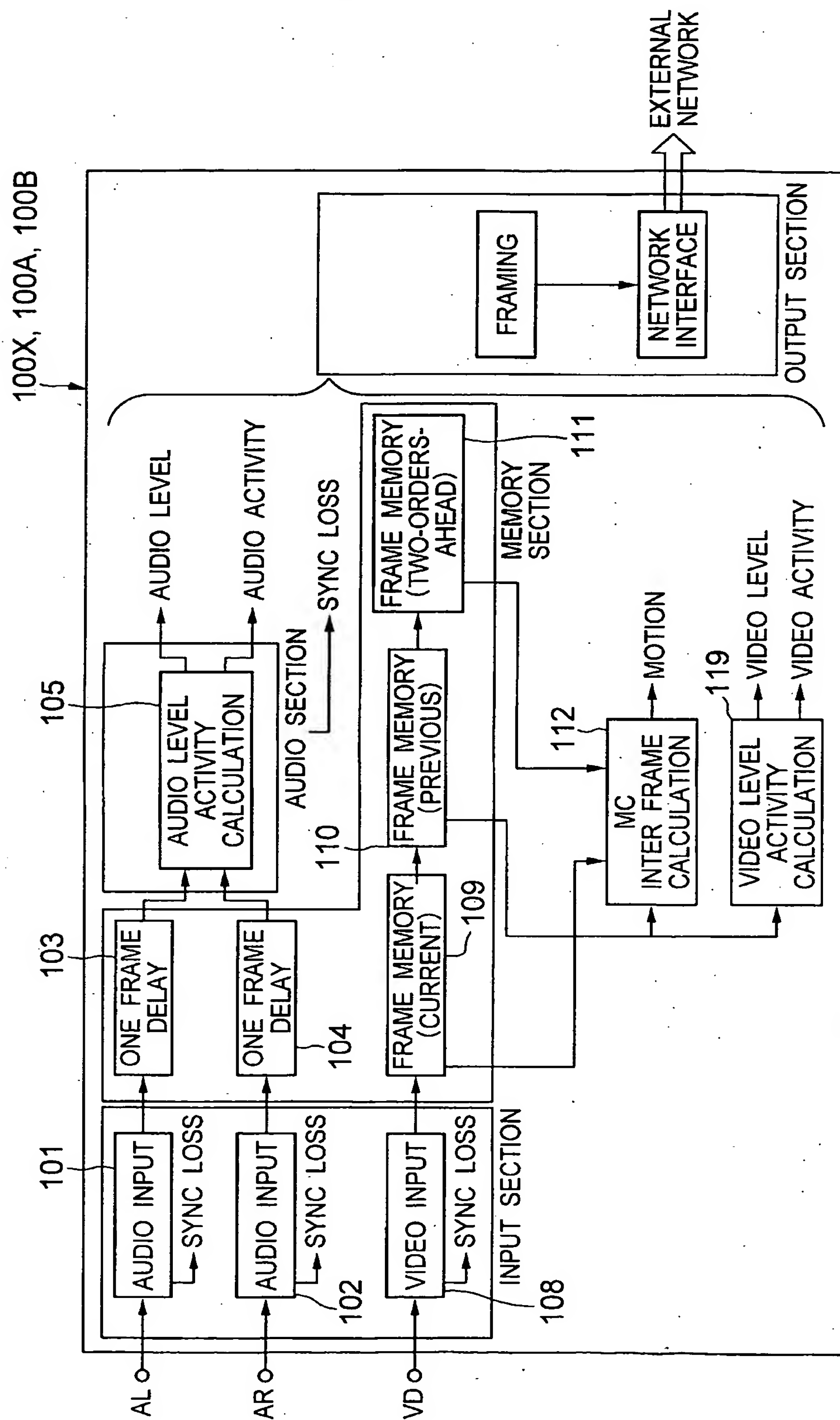


FIG. 2



3/8

FIG. 3A

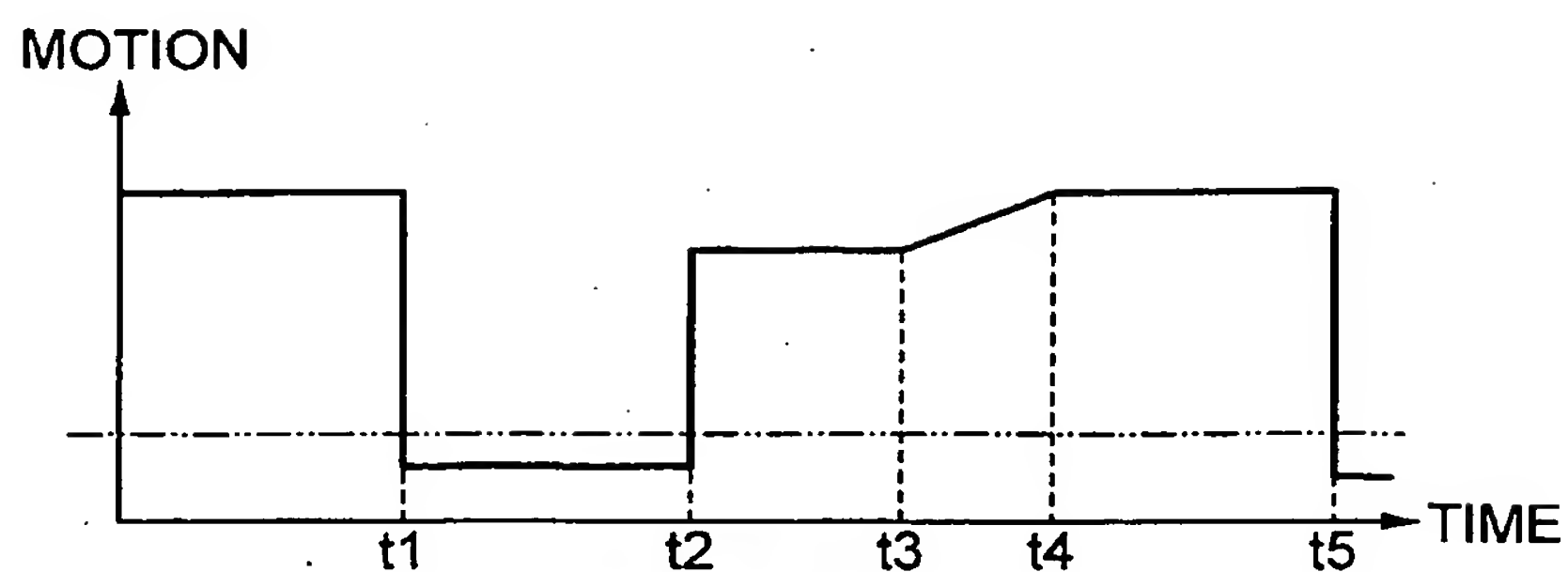


FIG. 3B

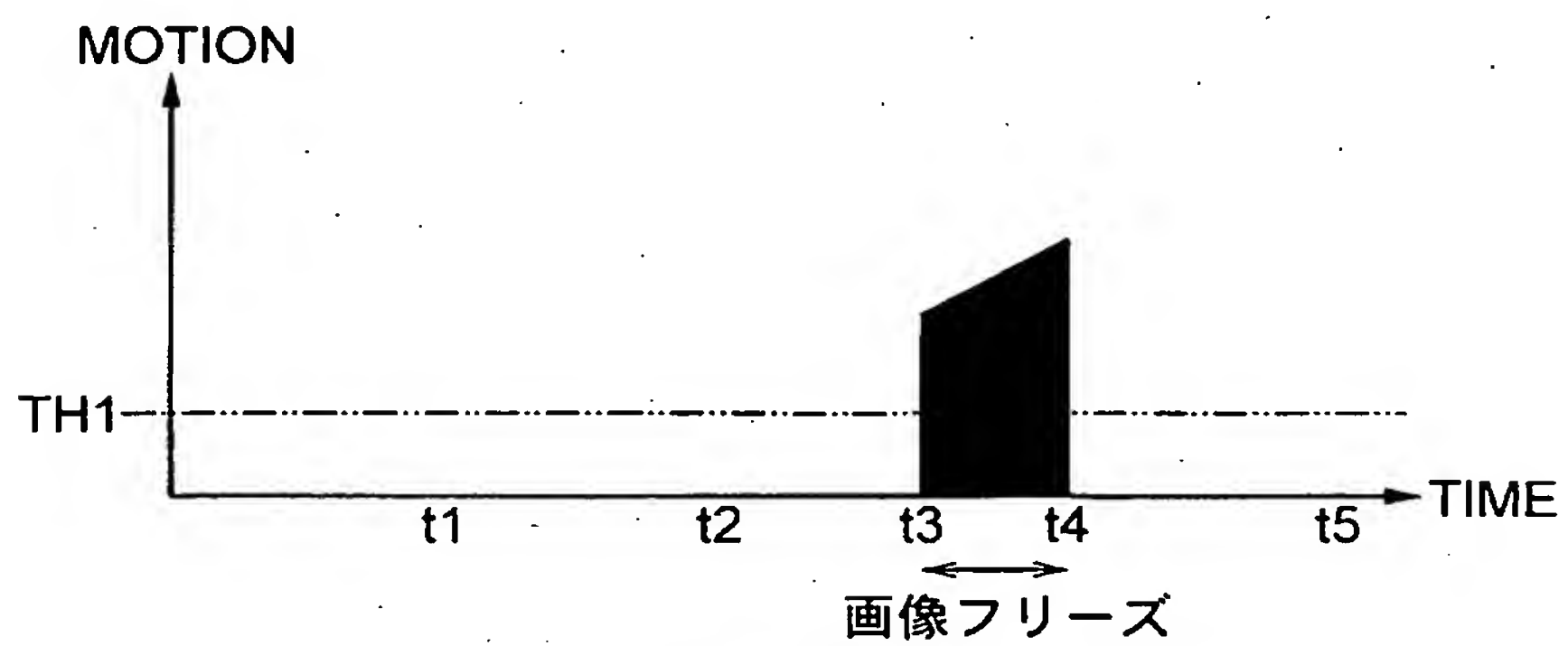
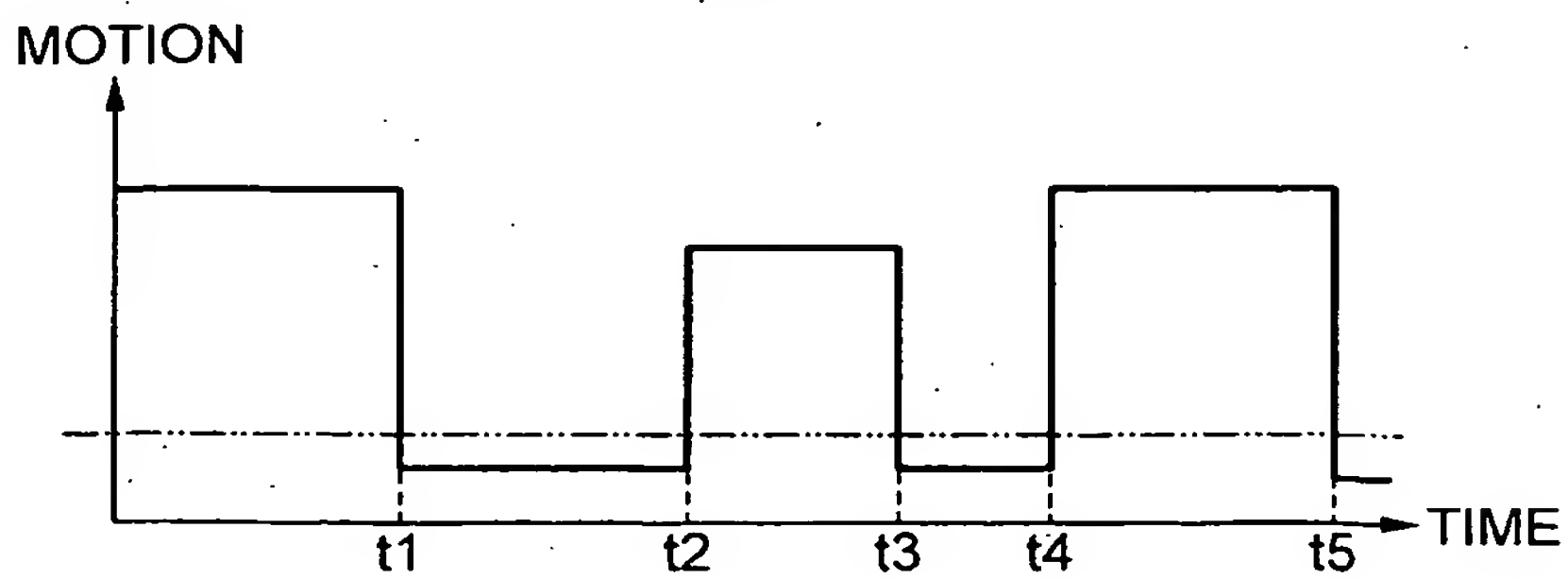


FIG. 3C



4/8

FIG. 4A

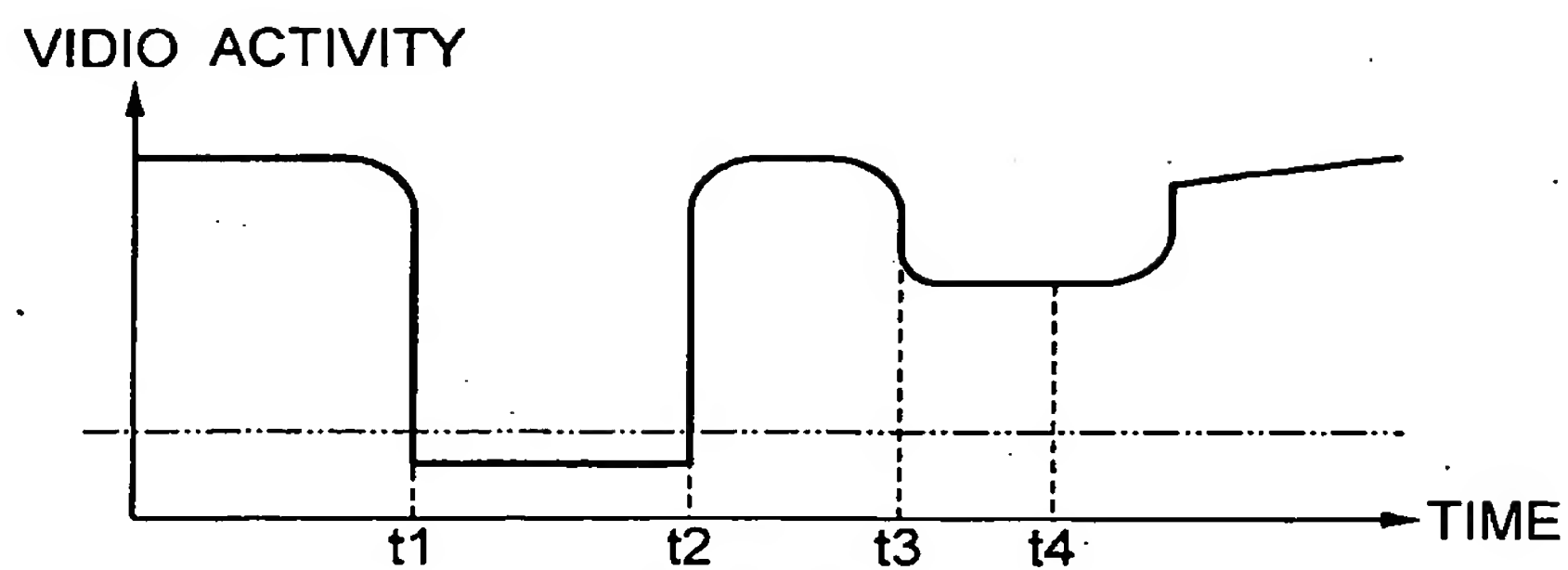


FIG. 4B

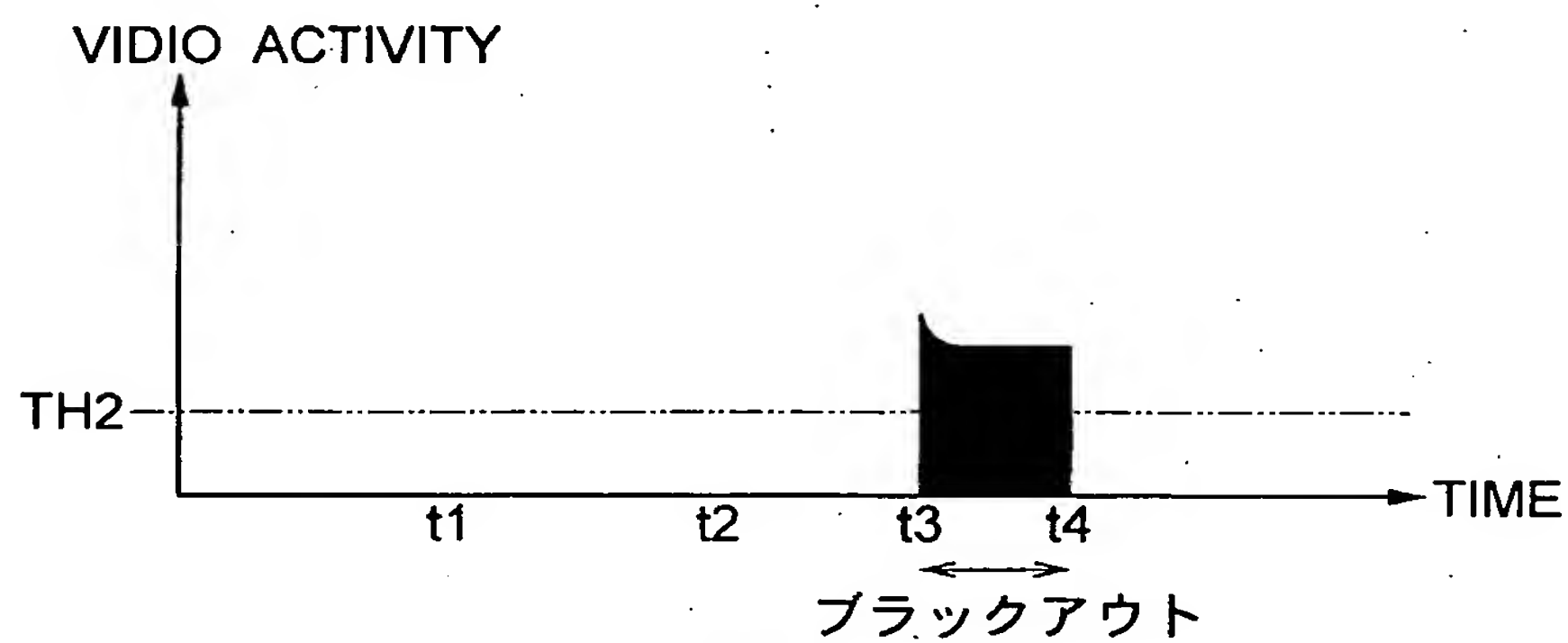
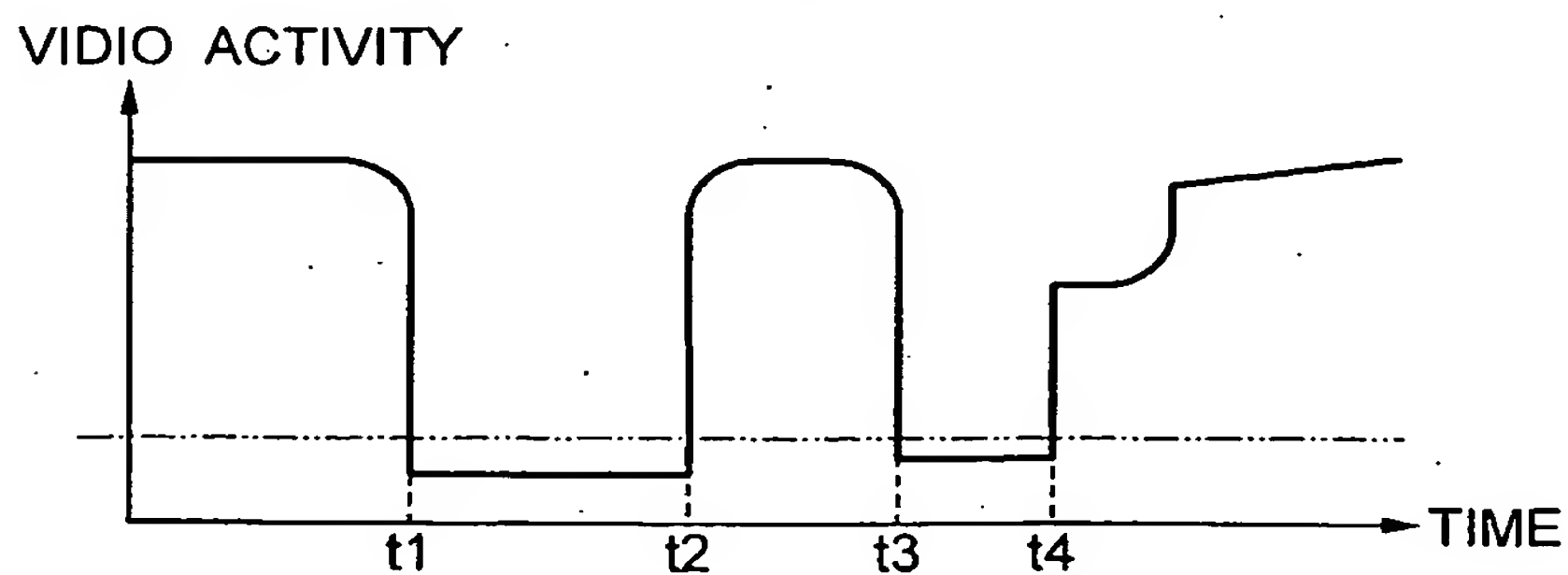


FIG. 4C



5/8

FIG. 5A

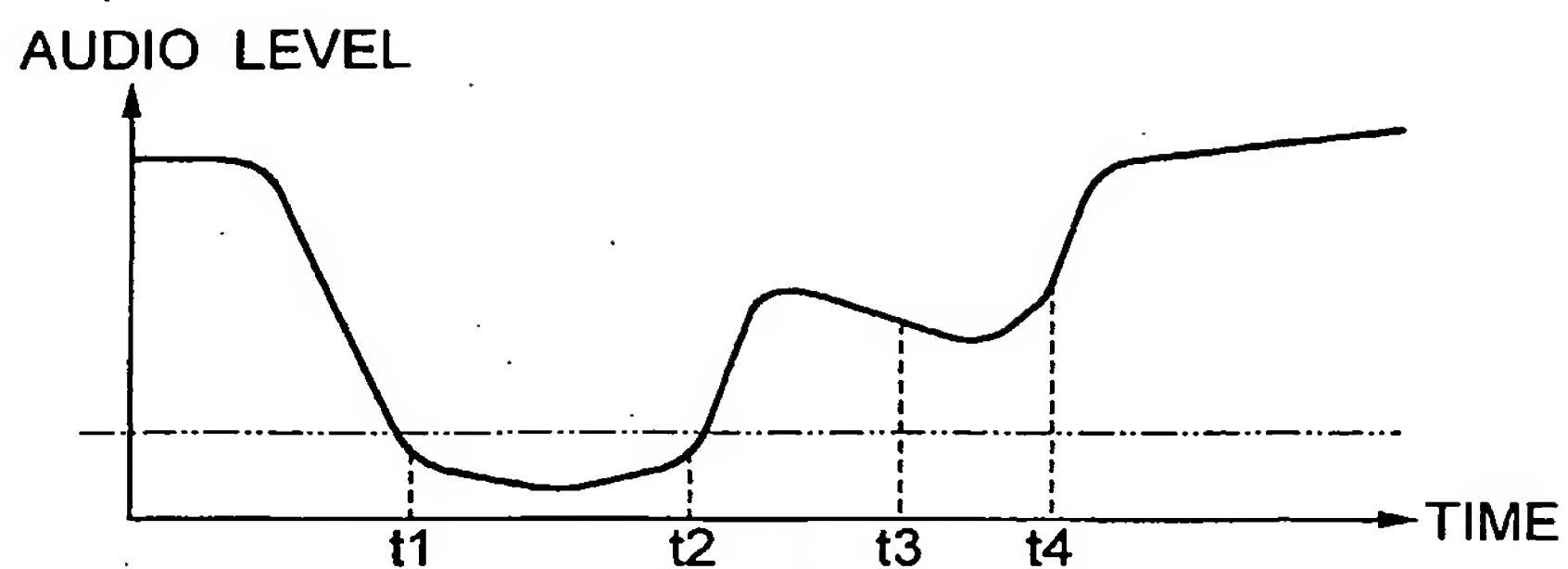


FIG. 5B

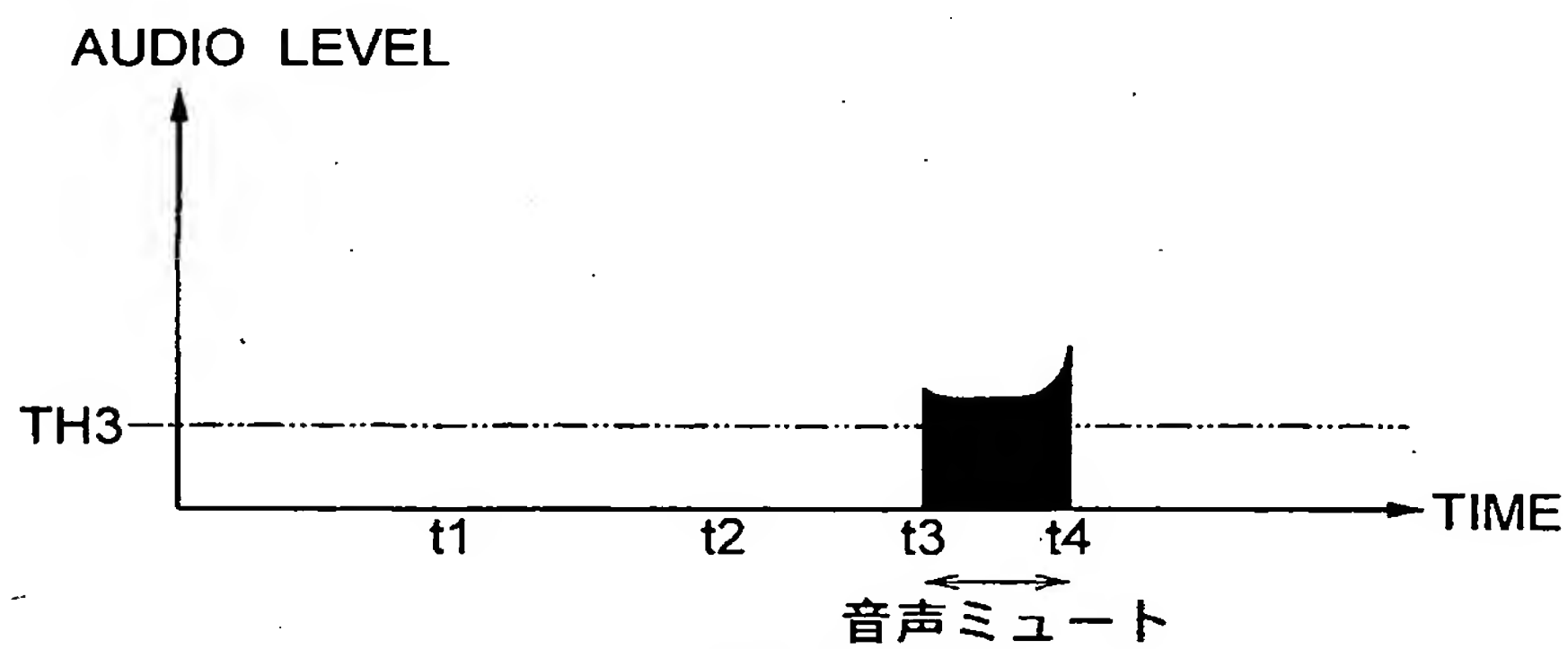
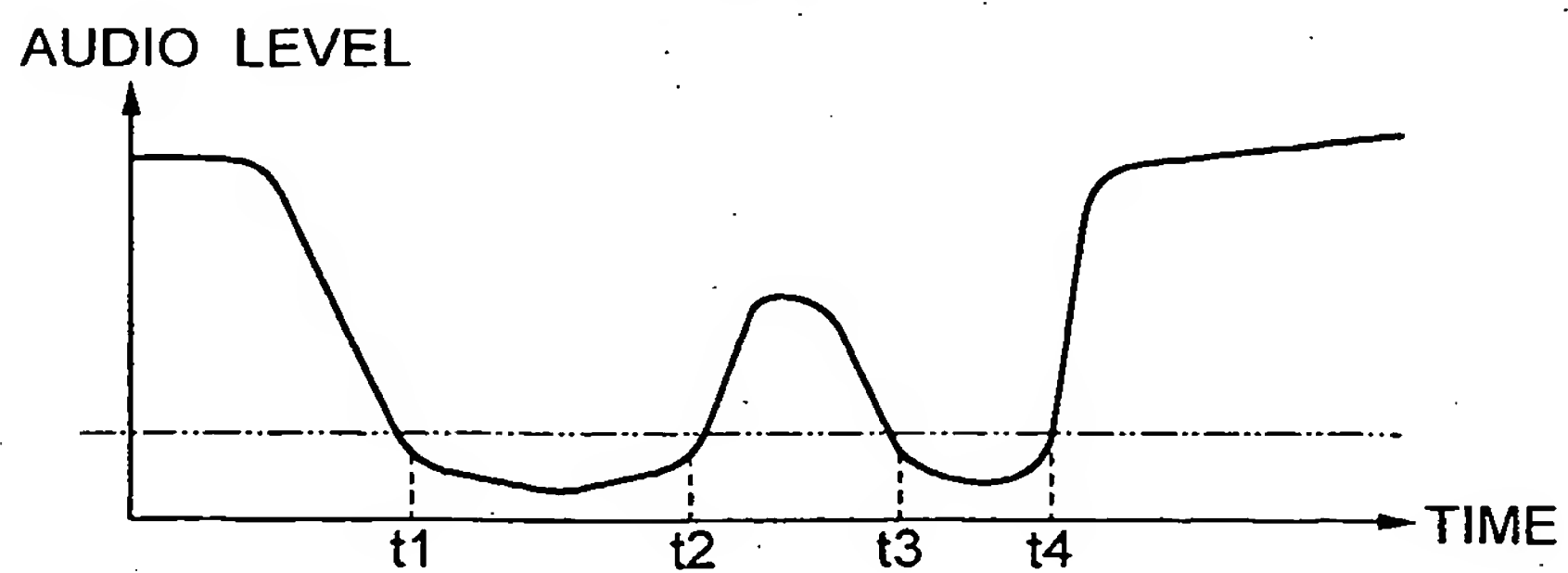


FIG. 5C



6/8

FIG. 6A

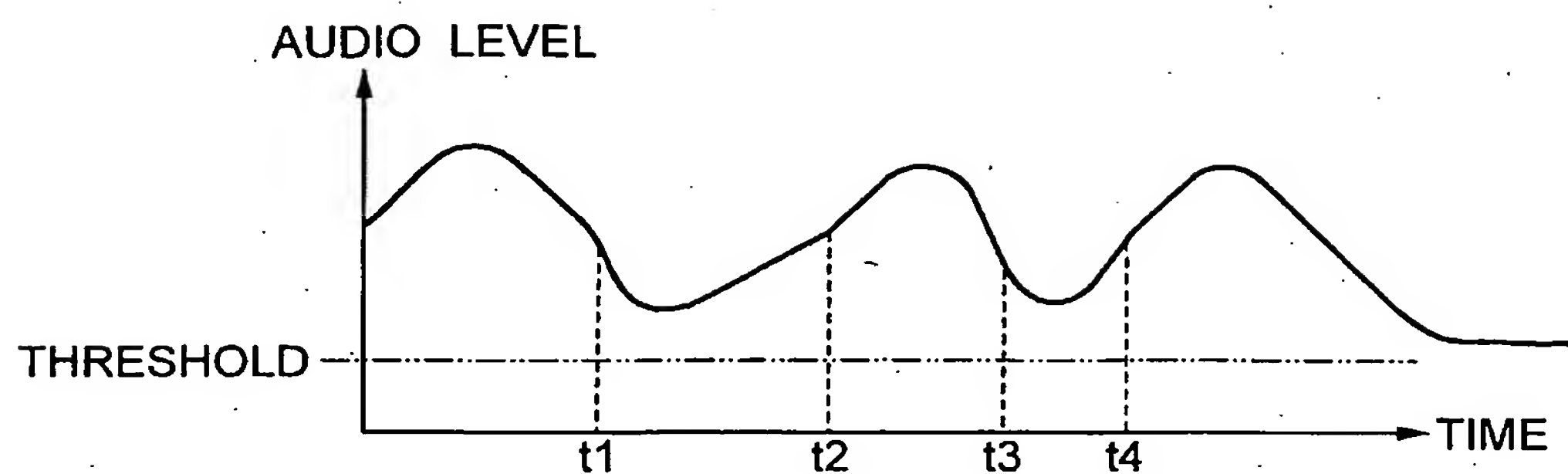


FIG. 6B

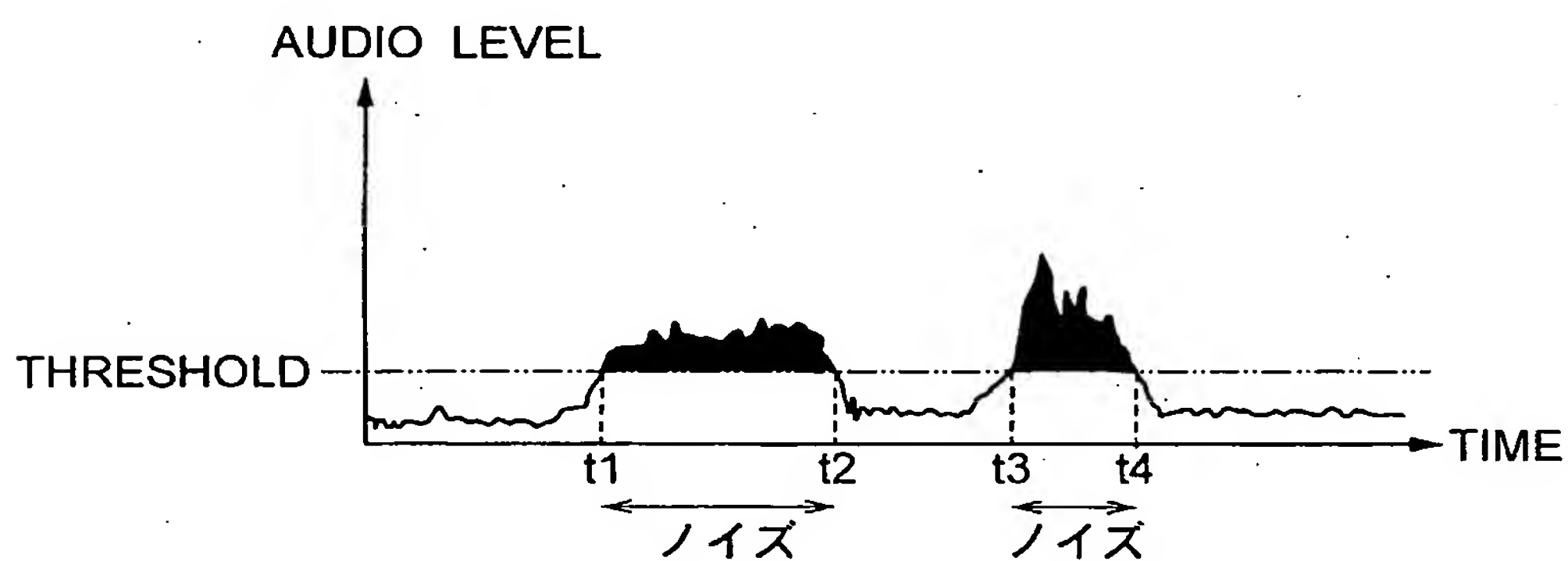


FIG. 6C

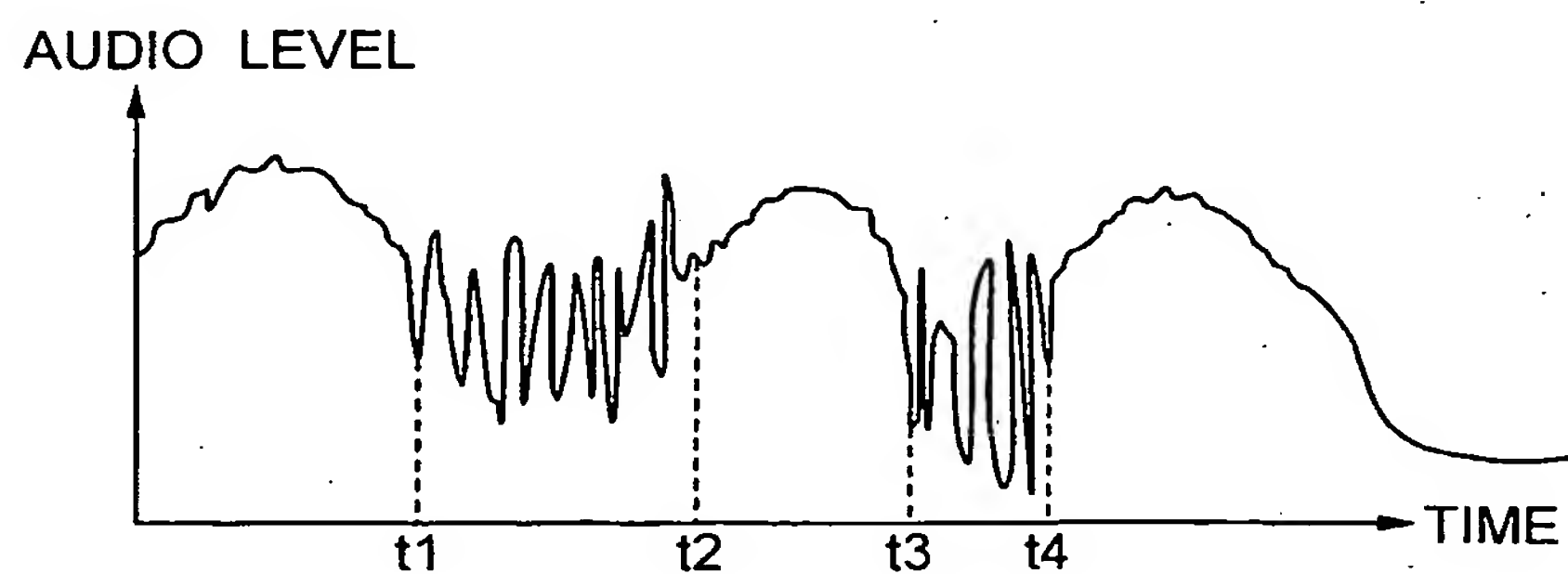


FIG. 7A

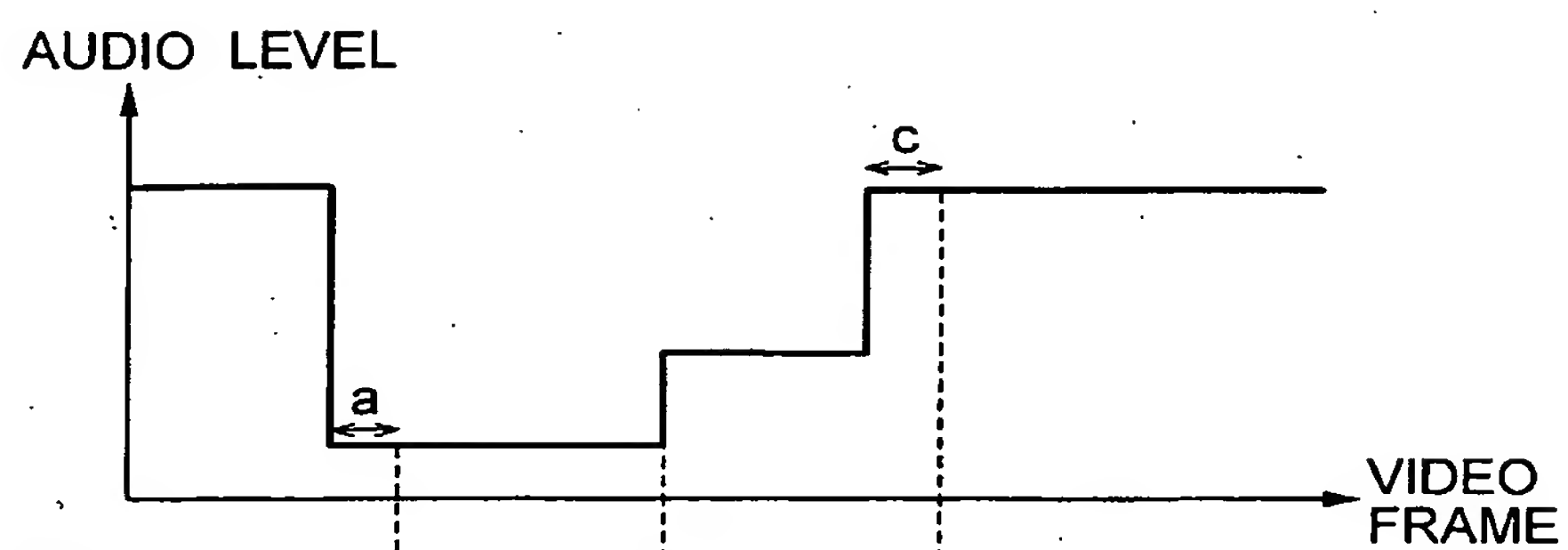


FIG. 7B

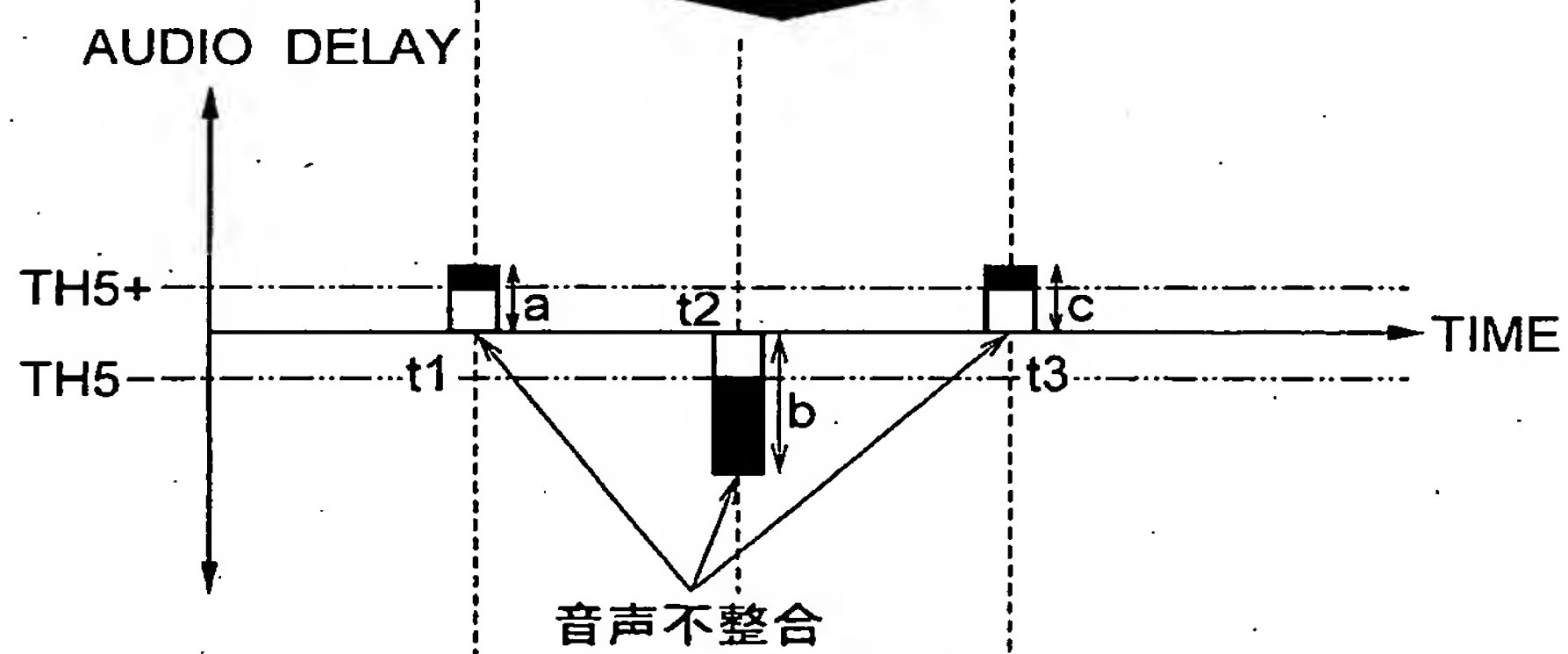


FIG. 7C

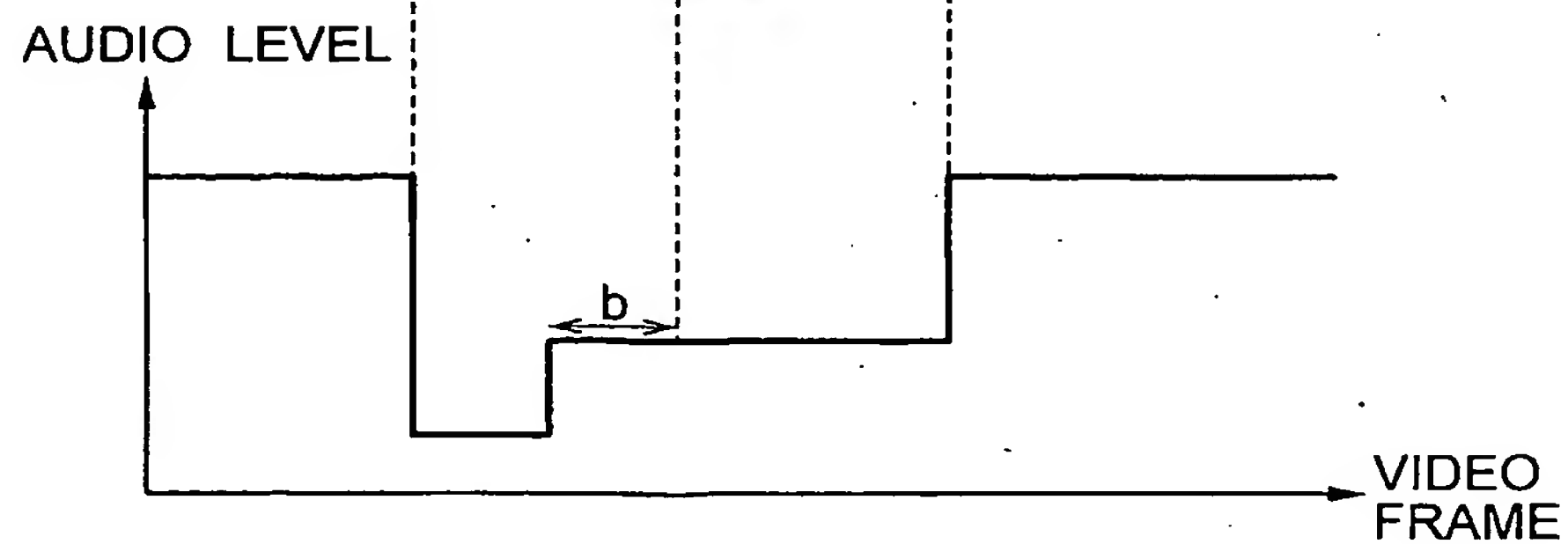


FIG. 8A

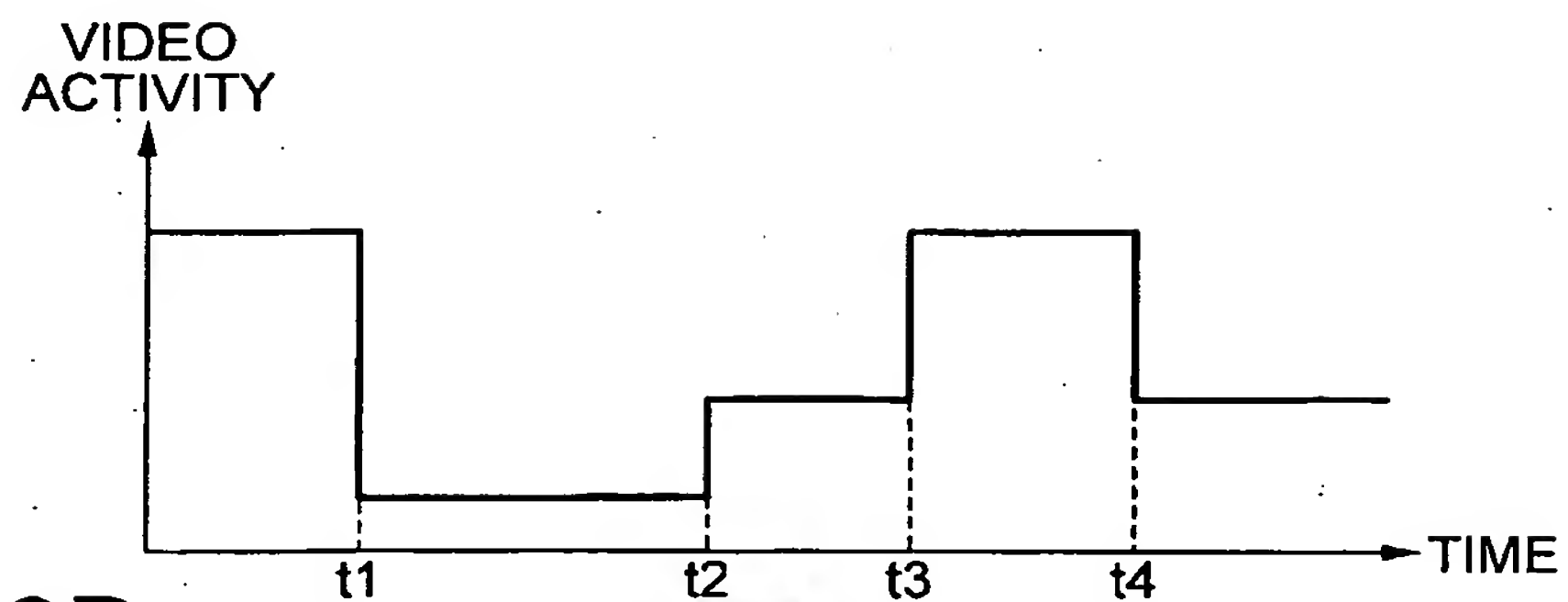


FIG. 8B

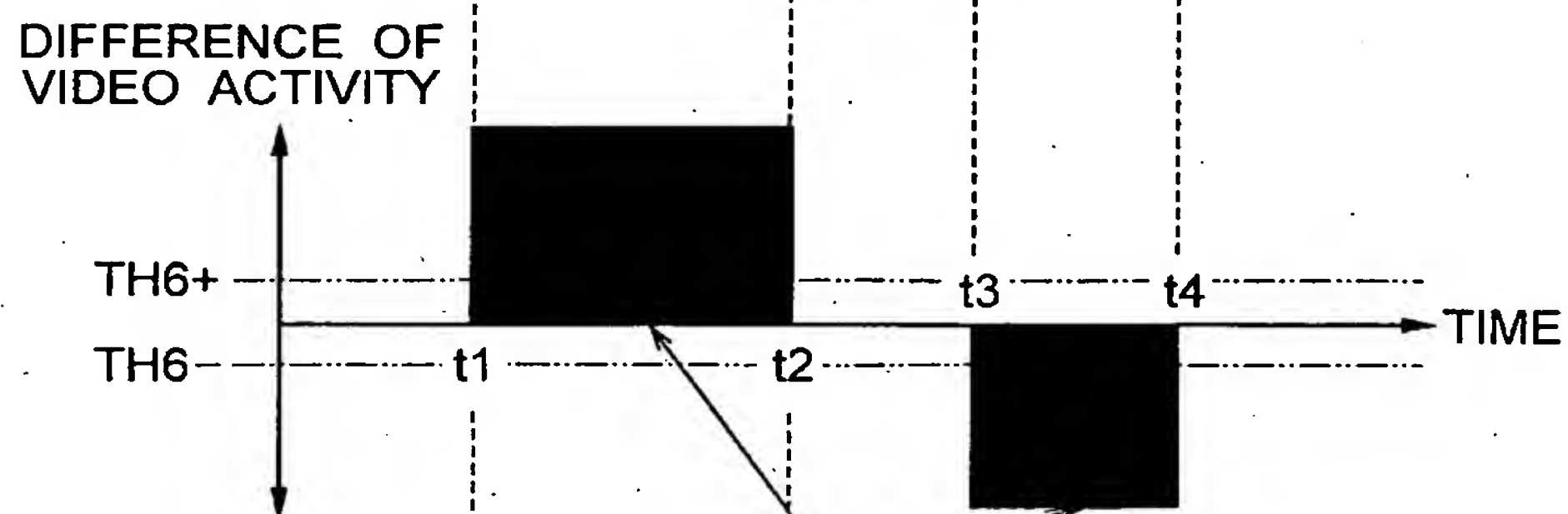


FIG. 8C

